

An averaging principle for diffusions in foliated spaces

Iván I. Gonzáles Gargate¹

Paulo R. Ruffino²

*Departamento de Matemática, Universidade Estadual de Campinas,
13.083-859- Campinas - SP, Brazil.*

Abstract

Consider an SDE on a foliated manifold whose trajectories lay on compact leaves. We investigate the effective behaviour of a small transversal perturbation of order ϵ . An average principle is shown to hold such that the component transversal to the leaves converges to the solution of a deterministic ODE, according to the average of the perturbing vector field with respect to invariant measures on the leaves, as ϵ goes to zero. An estimate of the rate of convergence is given. These results generalize the geometrical scope of previous approaches, including completely integrable stochastic Hamiltonian system.

Key words: Averaging principle, foliated diffusion, rescaled stochastic systems, stochastic flows.

MSC2010 subject classification: 60H10, 58J65, 58J37.

1 Introduction and set up

Generally speaking, the original heuristic idea of an averaging principles refers to an intertwining of two dynamics where one of them is, in some sense, much slower and is affected somehow by the other faster dynamics. An averaging principle in this case refers to the possibility of approximate, in some topology, the slow dynamics considering only the average action or perturbation which the fast motion induces on it. These ideas have appeared long ago and, as mentioned by V. Arnold [3, p.287], they were implicitly contained in the works of Laplace, Lagrange and Gauss on celestial mechanics; literature on the matter can be found e.g. among many others in [3], Sanders, Verhulst and Murdoch [13] and references therein. Presently, on what regards stochastic systems, averaging has been quite an active research field on which there is also a vast literature on the topic. Interesting quick historical overviews can be found in X.-M. Li [9, p.806], Kabanov and Pergamenshchikov [7, Appendix], [13, Appendix A]. Among many other works somehow related to the topic, we refer to Khasminski and Krylov [8], Sowers [14], Namachchaya and Sowers [10] Borodin and Freidlin [4], [7] and references therein.

The specific problem that we address in this article is a perturbation of a diffusion in a foliated manifold M such that the unperturbed random trajectories lay on the leaves. The perturbation are taken transversal to the leaves of the foliation. Here, the slow system is the transversal component and the fast system is given by the

¹e-mail: ivan.gargate@unila.edu.br.

²Corresponding author, e-mail: ruffino@ime.unicamp.br.

rescaled y_t^ϵ , where y_t^ϵ is the solution of the original SDE perturbed by a vector field ϵK .

Our results generalize the recent approach by X.M. Li [9] on an averaging principle for a completely integrable stochastic Hamiltonian system. In that article, as in the classical approach, see e.g. [3], Li has explored the benefits of a well structured geometrical coordinates in the state space given by the coordinates of the Liouville torus; these benefits include vanishing Itô-Stratonovich correction terms besides also vanishing covariant derivative of Hamiltonian vector fields in tangent directions to the leaves. We prove here that an averaging principle also holds in a generalized geometrical scope, so that this averaging phenomenon occurs independently of symplectic structures (but with possibly slower rates of convergence). Comparing to Li's previous result [9, Lemma 3.2], where the estimates contain a term of order $1/\sqrt{t}$, our corresponding estimates in Lemma 3.1 are continuous at $\epsilon = t = 0$. Some of the rates of convergence of [9, Lemma 3.1] is recovered as particular cases in Corollaries 2.2 and 2.3. In the main result, we show that in the average, the approximation goes to zero faster than $|\ln \epsilon|^{-\frac{\beta}{p}}$, for $\beta \in (0, 1/2)$.

The set up. Let M be a smooth Riemannian manifold with an n -dimensional smooth foliation, i.e. M is endowed with an integrable regular distribution of dimension n (for definition and further properties of foliated spaces see e.g. the initial chapters of Tondeur [15], Walczak [16] among others). We denote by L_x the leaf of the foliation passing through a point $x \in M$. For simplicity, we shall assume that the leaves are compact and that each leaf L_x has a tubular neighbourhood $U \subset M$ where U is diffeomorphic to $L_x \times V$, where $V \subset \mathbf{R}^d$ is an open bounded neighbourhood of the origin and d is the codimension of the foliation. We shall assume an SDE in M whose solution flow preserves the foliation, i.e. we consider a Stratonovich equation

$$dx_t = X_0(x_t)dt + \sum_{i=1}^r X_i(x_t) \circ dB_t^i \quad (1)$$

where the smooth vector fields X_i are foliated in the sense that $X_i(x) \in T_x L_x$, for $i = 0, 1, \dots, r$. Here $B_t = (B_t^1, \dots, B_t^r)$ is a standard Brownian motion in \mathbf{R}^r with respect to a filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbf{P})$. For an initial condition x_0 , the trajectories of the solution x_t in this case lay on the leaf L_{x_0} a.s.. Moreover, there exists a (local) stochastic flow of diffeomorphisms $F_t : M \rightarrow M$ which restricted to the initial leaf is a flow in the compact L_{x_0} .

For a smooth vector field K in M , we shall denote the perturbed system by y_t^ϵ which satisfies the SDE

$$dy_t^\epsilon = X_0(y_t^\epsilon)dt + \sum_{i=1}^r X_i(y_t^\epsilon) \circ dB_t^i + \epsilon K(y_t^\epsilon) dt, \quad (2)$$

with the same initial condition $y_0^\epsilon = x_0$ of the unperturbed system x_t .

Our main result, Theorem 4.1, says that locally the transversal behaviour of y_t^ϵ can be approximated in the average by an ordinary differential equation in the

transversal space whose coefficients are given by the average of the transversal component of the perturbation K with respect to the invariant measure on the leaves for the original dynamics of Equation (1). The reader will notice by the end of the proofs that compactness of the leaves in fact can be substituted by some other boundedness conditions, added also to some rather technical adjustments which we will not address here. In the Sections 2 and 3 we present the main lemmas. The main result appears in Section 4, where we also present a simple illustrative example. In particular, under some symmetry hypothesis on a foliated system embedded in an Euclidean space, we use the main theorem to conclude that Lyapunov exponents in the transversal direction must tend to zero as ϵ goes to zero, cf. Proposition 4.2.

2 Preliminaries results

Our coordinate system. Given an initial condition $x_0 \in M$, let $U \subset M$ be a bounded neighborhood of x_0 which is diffeomorphic to $L_{x_0} \times V$ and whose closure $\bar{U} \subset M$. By compactness of L_{x_0} , there exists a finite number of local foliated coordinate systems $\varphi_i : U_i \rightarrow W_i \times V \subset \mathbf{R}^n \times \mathbf{R}^d$, where W_i and V are open sets, say with $1 \leq i \leq k$ and $x_0 \in U_1$ such that:

- 1) $U = \cup_{i=1}^k U_i$;
- 2) The leaf $L_{x_0} = \cup_{i=1}^k \varphi_i^{-1}(W_i \times \{0\})$, i.e. each U_i is diffeomorphic to the product of an open set (with the induced topology) in the leaf L_{x_0} and the vertical component V ;
- 3) If a pair of points $p \in U_i$ and $q \in U_j$ in U belong to the same leaf then their transversal coordinates in V are the same; i.e. $\pi(\varphi_i(p)) = \pi(\varphi_j(p))$ where π is the projection on the transversal space V ;
- 4) For $i = 1, \dots, k$, φ_i has bounded derivatives (obtained reducing open set U if necessary).

Note that for a fixed $y \in V$, the finite union $\cup_i \varphi_i^{-1}(W_i, y)$ is the leaf $L_{\varphi_i^{-1}(x, y)}$ for any $x \in W_i$. Natural examples of this scheme of coordinates systems appear if we consider compact foliation given by the inverse image of submersions: values in the image space provide local coordinates for the vertical space V .

Next lemma gives information on the order of which the perturbed trajectories y_t^ϵ approaches the unperturbed x_t when one varies ϵ and t in equation (2); it will be used to prove that the dynamics of the rescaled system y_t^ϵ is such that its time average for any function g in M approximates the time average of the spacial average of g on the leaves, Lemma 3.1. An exponential factor in the estimates is expected, as trivial linear examples show.

We shall denote the coordinates of a point $p \in U_1$ by $\varphi_1(p) = (u, v) \in \mathbf{R}^n \times \mathbf{R}^d$.

Lemma 2.1 *Let τ^ϵ be the first time the process y_t^ϵ exits the foliated coordinate neighbourhood U_1 as above. For any locally Lipschitz continuous function $f : M \rightarrow \mathbf{R}$ and $2 \leq p < \infty$ we have*

$$\left[\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |f(y_s^\epsilon) - f(x_s)|^p \right) \right]^{\frac{1}{p}} \leq K_1 \epsilon t e^{K_2 t^p}.$$

where $K_1, K_2 \geq 0$ are constants depending on upper bounds of the norms of the perturbing vector field K , on the Lipschitz coefficients of f and on the derivatives of X_0, X_1, \dots, X_r with respect to the coordinate system.

Proof: Initially write x_t and y_t^ϵ , the solutions of Equations (1) and (2) respectively, according to the foliated coordinates φ_1 . So we write $(u_t, v_t) := \varphi_1(x_t)$ and $(u_t^\epsilon, v_t^\epsilon) := \varphi_1(y_t^\epsilon)$. Then

$$\begin{aligned} |f(y_t^\epsilon) - f(x_t)| &= |f \circ \varphi_1^{-1}(u_t^\epsilon, v_t^\epsilon) - f \circ \varphi_1^{-1}(u_t, v_t)| \\ &\leq C |u_t^\epsilon - u_t| + C |v_t^\epsilon - v_t|, \end{aligned} \quad (3)$$

for some constant $C \geq 0$, using the fact that U_1 is relatively compact.

We shall denote $u_t = (u_t^1, \dots, u_t^n)$, $v_t = (v_t^1, \dots, v_t^d)$, $u_t^\epsilon = (u_t^{\epsilon,1}, \dots, u_t^{\epsilon,n})$, $v_t^\epsilon = (v_t^{\epsilon,1}, \dots, v_t^{\epsilon,d})$. We also split the horizontal and vertical component of the perturbing vector field $\tilde{K} = (K_u, K_v)$, into coordinates: $K_u = (K_u^1, \dots, K_u^n)$ and $K_v = (K_v^1, \dots, K_v^d)$.

In our coordinate system, the equations of the horizontal and vertical components of the perturbed system u_t^ϵ and v_t^ϵ are given by

$$du_t^{\epsilon,i} = \sum_{k=1}^r b_k^i(u_t^\epsilon, v_t^\epsilon) \circ dB_t^k + b_0^i(u_t^\epsilon, v_t^\epsilon) dt + \epsilon K_u^i(u_t^\epsilon, v_t^\epsilon) dt, \quad (4)$$

$$dv_t^{\epsilon,j} = \epsilon K_v^j(u_t^\epsilon, v_t^\epsilon) dt, \quad (5)$$

with $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, d$, for the induced vector fields b_0, b_1, \dots, b_r which, together with their derivatives, are also bounded. From equation (5) we have

$$\begin{aligned} \sup_{s \leq t \wedge \tau^\epsilon} |v_s^\epsilon - v_s| &\leq \epsilon \sup_{s \leq t \wedge \tau^\epsilon} \int_0^s |K_v(u_s^\epsilon, v_s^\epsilon)| ds \\ &\leq \epsilon t \sup_{x \in U} |K_v(x)| = C_1 \epsilon t, \end{aligned} \quad (6)$$

where $C_1 = \sup_{x \in U} |K(x)|$. From equation (4) we have in each i -th component, for $s < \tau^\epsilon$:

$$u_s^{\epsilon,i} - u_s^i = \sum_{k=1}^r \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) \circ dB_r^k + \quad (7)$$

$$\int_0^s (b_0^i(u_r^\epsilon, v_r^\epsilon) - b_0^i(u_r, v_r)) dr + \epsilon \int_0^s K_u^i(u_r^\epsilon, v_r^\epsilon) dr. \quad (8)$$

In terms of Itô integral,

$$\begin{aligned} \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) \circ dB_r^k &= \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \\ &\quad + \frac{1}{2} \int_0^s [\nabla b_k^i \cdot b_k(u_r^\epsilon, v_r^\epsilon) - \nabla b_k^i \cdot b_k(u_r, v_r)] dr. \end{aligned}$$

Hence, taking the absolute values in both sides of Equation (8) we get, for each i :

$$\begin{aligned} |u_s^{\epsilon, v} - u_s^i| &\leq \sum_{k=1}^r \left| \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right| + \\ &\quad \frac{1}{2} \sum_{k=1}^r \int_0^s |\nabla b_k^i \cdot b_k(u_r^\epsilon, v_r^\epsilon) - \nabla b_k^i \cdot b_k(u_r, v_r)| dr \\ &\quad + \int_0^s |b_0^i(u_r^\epsilon, v_r^\epsilon) - b_0^i(u_r, v_r)| dr + \epsilon \int_0^s |K_u^i(u_r^\epsilon, v_r^\epsilon)| dr. \end{aligned} \quad (9)$$

Functions b_0^i and $(\nabla b_k^i \cdot b_k)$ are Lipschitz, hence for a common constant C_2 ,

$$\begin{aligned} |u_s^{\epsilon, i} - u_s^i| &\leq \left| \sum_{k=1}^r \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right| + \\ &\quad C_2 \int_0^s |v_r^\epsilon - v_r| dr + C_2 \int_0^s |u_r^\epsilon - u_r| dr + \epsilon s \sup_U |K_u|. \end{aligned} \quad (10)$$

The first deterministic integral, together with inequality (6) yields:

$$\begin{aligned} |u_s^{\epsilon, i} - u_s^i| &\leq \sum_{k=1}^r \left| \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right| + C_1 C_2 \epsilon s^2 \\ &\quad + C_2 \int_0^s |u_r^\epsilon - u_r| dr + C_1 \epsilon s. \end{aligned}$$

Now, for $p \geq 1$, there exists a constant C_3 such that

$$\begin{aligned} |u_s^{\epsilon, i} - u_s^i|^p &\leq C_3 \sum_{k=1}^r \left| \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right|^p + C_3 (C_1 C_2 \epsilon s^2)^p \\ &\quad + C_3 C_2^p \left(\int_0^s |u_r^\epsilon - u_r| dr \right)^p + C_3 (C_1 \epsilon s)^p. \end{aligned}$$

Cauchy-Schwartz inequality yields:

$$\begin{aligned} |u_s^{\epsilon, i} - u_s^i|^p &\leq C_3 \sum_{k=1}^r \left| \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right|^p + C_3 (C_1 C_2 \epsilon s^2)^p \\ &\quad + C_3 C_2^p s^{p-1} \int_0^s |u_r^\epsilon - u_r|^p dr + C_3 (C_1 \epsilon s)^p. \end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^{\epsilon,i} - u_s^i|^p &\leq C_3 \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} \left| \sum_{k=1}^r \int_0^s (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r)) dB_r^k \right|^p + C_3 (C_1 C_2 \epsilon t^2)^p \\
&\quad + C_3 C_2^p t^{p-1} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} \int_0^s |u_r^\epsilon - u_r|^p dr + C_3 (C_1 \epsilon t)^p \\
&\leq C_4 \sum_{k=1}^r \mathbb{E} \left[\int_0^{t \wedge \tau^\epsilon} (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r))^2 dr \right]^{p/2} + C_3 (C_1 C_2 \epsilon t^2)^p \\
&\quad + C_3 C_2^p t^{p-1} \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr + C_3 (C_1 \epsilon t)^p
\end{aligned}$$

where we have used classical L^p -inequality for martingales (e.g. Revuz and Yor [11]). Using again the Lipchitz property of each b_k for the terms in the brackets above:

$$\begin{aligned}
\sum_{k=1}^r \int_0^{t \wedge \tau^\epsilon} (b_k^i(u_r^\epsilon, v_r^\epsilon) - b_k^i(u_r, v_r))^2 dr &\leq 2C_2^2 \left(\int_0^{t \wedge \tau^\epsilon} |v_r^\epsilon - v_r|^2 dr + \int_0^{t \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^2 dr \right) \\
&\leq 2C_2^2 \left(\int_0^t C_1^2 \epsilon^2 r^2 dr + \int_0^{t \wedge \tau^\epsilon} \sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^2 dr \right) \\
&\leq C_2^2 C_1^2 \epsilon^2 t^3 + 2C_2^2 \int_0^{t \wedge \tau^\epsilon} \sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^2 dr. \tag{11}
\end{aligned}$$

We end up with:

$$\begin{aligned}
\mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^{\epsilon,i} - u_s^i|^p &\leq C_4 \sum_{k=1}^r \mathbb{E} \left[C_2^2 C_1^2 \epsilon^2 t^3 + 2C_2^2 \int_0^{t \wedge \tau^\epsilon} \sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^2 dr \right]^{p/2} \\
&\quad + C_3 C_2^p t^{p-1} \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr \\
&\quad + C_3 (C_1 C_2 \epsilon t^2)^p + C_3 (C_1 \epsilon t)^p. \tag{12}
\end{aligned}$$

For $p \geq 2$ one can use Cauchy-Schwartz again to conclude that there exists a constant C_5 such that the last expression is less than or equal

$$\begin{aligned}
&C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_5 C_2^p t^{\frac{p-2}{2}} \int_0^{t \wedge \tau^\epsilon} \mathbb{E} \sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p dr + C_3 (C_1 C_2 \epsilon t^2)^p \\
&+ C_3 C_2^p t^{p-1} \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr + C_3 (C_1 \epsilon t)^p
\end{aligned}$$

$$\begin{aligned}
&= C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_3 \left(C_1 C_2 \epsilon t^2 \right)^p + C_3 \left(C_1 \epsilon t \right)^p \\
&\quad + \left(C_5 C_2^p t^{\frac{p-2}{2}} + C_3 C_2^p t^{p-1} \right) \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr.
\end{aligned}$$

Now, summing up over i in the inequalities above leads to

$$\begin{aligned}
\mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p &\leq C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_3 \left(C_1 C_2 \epsilon t^2 \right)^p + C_3 \left(C_1 \epsilon t \right)^p \\
&\quad + \left(C_5 C_2^p t^{\frac{p-2}{2}} + C_3 C_2^p t^{p-1} \right) \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr.
\end{aligned}$$

We use now the integral form of Gronwall's inequality to find that:

$$\begin{aligned}
\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) &\leq C_6 \epsilon^p t^p (1 + t^p) \exp\{C_7(t^{p/2} + t^p)\} \\
&\leq C_8 \epsilon^p t^p (1 + t^p) \exp\{C_9 t^p\}.
\end{aligned}$$

Going back to the inequality 3, now we have

$$|f(y_t^\epsilon) - f(x_t)|^p \leq C_{10} |v_t^\epsilon - v_t|^p + C_{10} |u_t^\epsilon - u_t|^p$$

hence:

$$\begin{aligned}
\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |f(y_s^\epsilon) - f(x_s)|^p \right) &\leq C_{10} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |v_s^\epsilon - v_s|^p + C_{10} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^\epsilon - u_s|^p \\
&\leq C_{11} \epsilon^p t^p + C_{12} \epsilon^p t^p (1 + t^p) \exp(C_9 t^p) \\
&\leq C_{13} \epsilon^p t^p (1 + t^p) \exp(C_9 t^p).
\end{aligned}$$

From here, finally, one concludes that there exist constants K_1 and K_2 such that

$$\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |f(y_s^\epsilon) - f(x_s)|^p \right)^{\frac{1}{p}} \leq K_1 \epsilon t \exp(K_2 t^p).$$

□

Next corollary includes the case of a completely integrable stochastic Hamiltonian system when one uses the action-angle coordinates, cf. X.-M. Li [9, Lemma 3.1]).

Corollary 2.2 *If the vector fields X_0, \dots, X_r depend only on the vertical coordinate (null derivative in the directions of the leaves, as in the Hamiltonian case [9]) then the estimates above can be improved, and for $p \geq 1$ there exists a constant K_1 such that*

$$\left[\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |f(y_s^\epsilon) - f(x_s)|^p \right) \right]^{\frac{1}{p}} \leq K_1 \epsilon (t + t^2).$$

Proof: In this case the correction term of the Stratonovich stochastic integral in terms of Itô integral in inequality (9) vanishes, and also so does the determinist integration of $|u_r^\epsilon - u_r|$ in inequalities (10) and (11). Hence inequality (12) improves to

$$\mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^{\epsilon,i} - u_s^i|^p \leq C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_3 (C_1 C_2 \epsilon t^2)^p + C_3 (C_1 \epsilon t)^p. \quad (13)$$

The argument in the rest of the proof follows straightforward for $p \geq 1$ skipping Gronwall inequality. \square

Next corollary includes the case $X_0 \equiv 0$, cf. [9, Lemma 3.1(2)] for stochastic Hamiltonian systems with action-angle coordinate system.

Corollary 2.3 *If in addition to conditions of Corollary 2.2 above, we have that the deterministic vector field X_0 is constant when represented with respect to a certain local coordinate system in U_1 (i.e. b_0 has null derivative w.r.t u and v) then, for $p \geq 1$ the estimates can be improved further to $K_1 \epsilon (t + t^{\frac{3}{2}})$.*

Proof: Besides the vanishing terms already mentioned above, the second deterministic integral on the right hand side of inequality (9) also vanishes. Hence inequality (12) simplifies further to

$$\mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^{\epsilon,i} - u_s^i|^p \leq C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_3 (C_1 \epsilon t)^p.$$

\square

Yet, from the proof of the Lemma 2.1 we have the following

Remark 2.4 *For $1 \leq p < 2$ and t sufficiently small, there exist constants K_1 and K_2 such that*

$$\left[\mathbb{E} \left(\sup_{s \leq t \wedge \tau^\epsilon} |f(y_s^\epsilon) - f(x_s)|^p \right) \right]^{\frac{1}{p}} \leq K_1 \epsilon t \exp(K_2 t^p). \quad (14)$$

Proof: One can no longer use Cauchy-Schwartz after inequality (12). Alternatively, from (12), use that

$$\begin{aligned} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^\epsilon - u_s|^p &\leq C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + C_5 C_2^p t^{p/2} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^{\epsilon,i} - u_s^i|^p + C_3 (C_1 C_2 \epsilon t^2)^p \\ &\quad + C_3 C_2^p \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr + C_3 C_1^p \epsilon^p t^p. \end{aligned}$$

If we fix an $0 < \delta < 1$ and take t sufficiently small such that $1 - C_5 C_2^p t^{p/2} > \delta$ then

$$\begin{aligned} \mathbb{E} \sup_{s \leq t \wedge \tau^\epsilon} |u_s^\epsilon - u_s|^p &\leq \delta^{-1} C_5 \left(C_2 C_1 \epsilon t^{3/2} \right)^p + \delta^{-1} C_3 (C_1 C_2 \epsilon t^2)^p \\ &\quad + \delta^{-1} C_3 C_2^p t^{p-1} \int_0^t \mathbb{E} \left(\sup_{s \leq r \wedge \tau^\epsilon} |u_r^\epsilon - u_r|^p \right) dr + \delta^{-1} C_3 C_1^p \epsilon^p t^p. \end{aligned}$$

And one completes the calculation as before using the integral version of Gronwall inequality. □

3 Averaging functions on the leaves

Consider a differentiable function $g : M \rightarrow \mathbb{R}$. The leaf L_p passing through a point $p \in M$ contains the support of an invariant measure μ_p for the unperturbed system (1); we shall assume that μ_p is ergodic. We shall work with the following function defined for each leaf $Q^g : V \subset \mathbf{R}^d \rightarrow \mathbf{R}$ given by the average of g with respect to these measures on the leaves. Namely, if v is the vertical coordinate of p , i.e. $p = \varphi(u, v)$, then:

$$Q^g(v) = \int_{L_p} g(x) d\mu_p(x).$$

We assume that Q^g has some degree of continuity with respect to v . This assumption appears in two levels in Lemma 3.4: 1) Riemann integrability of $Q^g(\pi(y_r^\epsilon))$ with respect to r guarantees the convergence to zero; 2) α -Hölder continuity guarantees the rate of convergence.

Our next step is to estimate the time average of g and Q^g along the perturbed system y_t^ϵ . To use the notations and results we have introduced before in local coordinates, we shall write simply $\pi(p) = v$ for the composition of the projection on the second coordinates with the local chart $\varphi(p) = (u, v)$.

Here the stopping time τ^ϵ denotes the first exit time of the open neighbourhood $U \subset M$ which is diffeomorphic to $L_{x_0} \times V$. We have the following estimates for the difference of the averages of functions g and Q^g .

Lemma 3.1 *Given a function $g : M \rightarrow \mathbf{R}$ let $Q^g : V \rightarrow \mathbf{R}$ be its average on the leaves. For $s, t \geq 0$ write*

$$\delta(\epsilon, t) = \int_{s \wedge \epsilon \tau^\epsilon}^{(s+t) \wedge \epsilon \tau^\epsilon} g(y_r^\epsilon) - Q^g(\pi(y_r^\epsilon)) dr.$$

Then $\delta(\epsilon, t)$ goes to zero when t or ϵ tend to zero.

Moreover, if Q^g is α -Hölder continuous with $\alpha > 0$ then for $p \geq 1$ and any $\beta \in (0, 1/2)$ we have the following estimates:

$$\left(\mathbb{E} \sup_{s \leq t} |\delta(\epsilon, s)|^p \right)^{\frac{1}{p}} \leq \sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}} h(t, \epsilon),$$

where $h(t, \epsilon)$ is continuous for $t, \epsilon > 0$ and converges to zero when $(t, \epsilon) \rightarrow 0$.

Proof:

The proof consists of considering a convenient partition of the interval $(s/\epsilon \wedge \tau^\epsilon, (s+t)/\epsilon \wedge \tau^\epsilon)$ where we can get the estimates by comparing in each subinterval the average of the flow of the original system (on the corresponding leaf) with the

average of the perturbed flow (possibly transversal to the leaves). These estimates in each subinterval are obtained using Lemma 2.1. So, a key point in the proof is a careful choice of the increments of such a convenient partition. For sufficiently small ϵ , we take the following assignment of increments:

$$\Delta t = \frac{(s+t) \wedge \tau^\epsilon - s \wedge \tau^\epsilon}{|\ln \epsilon|^{-\frac{2\beta}{p}}}.$$

Hence, the partition $t_n = \frac{s}{\epsilon} \wedge \tau^\epsilon + n\Delta t$, for $1 \leq n \leq N-1$, is such that

$$\frac{s}{\epsilon} \wedge \tau^\epsilon = t_0 < t_1 < \dots < t_{N-1} < \frac{s+t}{\epsilon} \wedge \tau^\epsilon.$$

with $N = N(\epsilon) = [\epsilon^{-1} |\ln \epsilon|^{-\frac{2\beta}{p}}]$ where here $[x]$ denotes the integer part of x .

Initially we represent the left hand side as the sum:

$$\epsilon \int_{\frac{s}{\epsilon} \wedge \tau^\epsilon}^{\frac{s+t}{\epsilon} \wedge \tau^\epsilon} g(y_r^\epsilon) dr = \epsilon \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} g(y_r^\epsilon) dr + \epsilon \int_{t_N}^{\frac{s+t}{\epsilon} \wedge \tau^\epsilon} g(y_r^\epsilon) dr.$$

Denote by θ_t the canonical shift operator on the probability space. Let $F_t(\cdot, \omega)$ with $t \geq 0$ be the flow of the original unperturbed system in M . Triangular inequality splits our calculation into four parts

$$|\delta(\epsilon, t)| \leq |A_1| + |A_2| + |A_3| + |A_4|, \quad (15)$$

where

$$\begin{aligned} A_1 &= \epsilon \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} [g(y_r^\epsilon) - g(F_{r-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))] dr, \\ A_2 &= \epsilon \sum_{n=0}^{N-1} \left[\int_{t_n}^{t_{n+1}} g(F_{r-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega))) dr - \Delta t Q^g(\pi(y_{t_n}^\epsilon)) \right], \\ A_3 &= \sum_{n=0}^{N-1} \epsilon \Delta t Q^g(\pi(y_{t_n}^\epsilon)) - \int_{s \wedge \epsilon \tau^\epsilon}^{(s+t) \wedge \epsilon \tau^\epsilon} Q^g(\pi(y_r^\epsilon)) dr, \\ A_4 &= \epsilon \int_{t_N}^{\frac{s+t}{\epsilon} \wedge \tau^\epsilon} g(y_r^\epsilon) dr. \end{aligned}$$

We proceed showing that each of the processes A_1, A_2, A_3 and A_4 above tends to zero uniformly on compact intervals. In what follows, we will explore many times the fact that for $a > 0$ and $b \in \mathbf{R}$, $\epsilon^a |\ln \epsilon|^b$ goes to zero as $\epsilon \searrow 0$. Hence, by construction, except when $\tau^\epsilon \leq s$ (where, restricted to which the lemma is trivial) both Δt and N go to infinity when ϵ tends to zero.

Lemma 3.2 *Process A_1 converges to zero uniformly on compact intervals when ϵ goes to zero. More precisely, we have the following estimates on the rate of convergence: For any $\gamma \in (0, 1)$, there exists a function h_1 such that*

$$\left(\mathbb{E} \sup_{s \leq t} |A_1|^p \right)^{\frac{1}{p}} \leq K_1 t \epsilon^\gamma h_1(t, \epsilon)$$

where h_1 is continuous in $t, \epsilon > 0$ and converges to zero when $(t, \epsilon) \mapsto (0, 0)$.

Proof: Initially note that by triangular inequality, and putting the supremum inside the integral we get

$$\left(\mathbb{E} \sup_{s \leq t} |A_1|^p \right)^{\frac{1}{p}} \leq \epsilon \sum_{n=0}^{N-1} \left(\mathbb{E} \left[\int_{t_n}^{t_{n+1}} \sup_{t_n \leq s \leq r} |g(y_s^\epsilon) - g(F_{s-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))| dr \right]^p \right)^{\frac{1}{p}}$$

If $\frac{1}{p} + \frac{1}{q} = 1$, by Hölder inequality we have that the estimate above is again bounded by

$$\epsilon \sum_{n=0}^{N-1} \left(\mathbb{E} \left[\left(\int_{t_n}^{t_{n+1}} dr \right)^{\frac{1}{q}} \left(\int_{t_n}^{t_{n+1}} \sup_{t_n \leq s \leq r} |g(y_s^\epsilon) - g(F_{s-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))|^p dr \right)^{\frac{1}{p}} \right]^p \right)^{\frac{1}{p}} \quad (16)$$

The increment Δt is a random variable bounded by $t |\ln \epsilon|^{\frac{2\beta}{p}}$. In the estimates below Δt will denote this very deterministic upper bounded. Analogously to N which has order $[\epsilon^{-1} |\ln \epsilon|^{-\frac{2\beta}{p}}]$. Hence, with this notation, last inequality is again bounded by

$$\begin{aligned} &\leq \epsilon (\Delta t)^{\frac{1}{q}} \sum_{n=0}^{N-1} \left(\mathbb{E} \left[\Delta t \sup_{t_n \leq s \leq t_{n+1}} |g(y_s^\epsilon) - g(F_{s-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))|^p \right] \right)^{\frac{1}{p}} \\ &\leq \epsilon \Delta t \sum_{n=0}^{N-1} \left(\mathbb{E} \left[\sup_{t_n \leq s \leq t_{n+1}} |g(y_s^\epsilon) - g(F_{s-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))|^p \right] \right)^{\frac{1}{p}} \end{aligned}$$

Lemma 2.1 and its corollaries says that for all $0 \leq n \leq N-1$ above, the function g evaluated along trajectories of the perturbed system compared with g evaluated along the unperturbed trajectories, both starting at $y_{t_n}^\epsilon$ satisfies:

$$\left[\mathbb{E} \sup_{t_n \leq s \leq t_{n+1}} |g(y_s^\epsilon) - g(F_{s-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega)))|^p \right]^{\frac{1}{p}} \leq K_1 \epsilon \Delta t e^{K_2 (\Delta t)^p}$$

Note that here, possibly Lemma 2.1 might be applied using different (finitely many) local coordinate systems φ_i for each n .

$$\left[\mathbb{E} \sup_{s \leq t} |A_1|^p \right]^{\frac{1}{p}} \leq K_1 N \epsilon^2 (\Delta t)^2 e^{K_2 (\Delta t)^p}$$

$$\begin{aligned}
&= K_1 \epsilon^2 [\epsilon^{-1} |\ln \epsilon|^{-\frac{2\beta}{p}}] t^2 |\ln \epsilon|^{\frac{4\beta}{p}} e^{K_2(t |\ln \epsilon|^{\frac{2\beta}{p}})^p} \\
&= K_1 t \epsilon^\gamma h_1(\epsilon, t)
\end{aligned}$$

for any $\gamma \in (0, 1)$ where

$$h_1(\epsilon, t) = t \epsilon^{\frac{1-\gamma}{2}} |\ln \epsilon|^{\frac{2\beta}{p}} \exp \left\{ \left(\frac{1-\gamma}{2} \right) \ln \epsilon + K_2 t^p |\ln \epsilon|^{2\beta} \right\}.$$

which satisfies the required properties for $\beta \in (0, 1/2)$. □

Lemma 3.3 *Process A_2 in equation (15) goes to zero with the following rate of convergence:*

$$\left[\mathbb{E} \sup_{s \leq t} |A_2|^p \right]^{\frac{1}{p}} \leq K \sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}}$$

for a positive constant K .

Proof: We have

$$\begin{aligned}
\left[\mathbb{E} \sup_{s \leq t} |A_2|^p \right]^{\frac{1}{p}} &\leq \epsilon \left[\mathbb{E} \left| \sum_{n=0}^{N-1} \left[\int_{t_n}^{t_{n+1}} g(F_{r-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega))) dr - \Delta t Q^g(\pi(y_{t_n}^\epsilon)) \right] \right|^p \right]^{\frac{1}{p}} \\
&\leq \epsilon \sum_{n=0}^{N-1} \left[\mathbb{E} \left| \int_{t_n}^{t_{n+1}} g(F_{r-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega))) dr - \Delta t Q^g(\pi(y_{t_n}^\epsilon)) \right|^p \right]^{\frac{1}{p}} \\
&= \epsilon \Delta t \sum_{n=0}^{N-1} \left[\mathbb{E} \left| \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} g(F_{r-t_n}(y_{t_n}^\epsilon, \theta_{t_n}(\omega))) dr - Q^g(\pi(y_{t_n}^\epsilon)) \right|^p \right]^{\frac{1}{p}}
\end{aligned}$$

For all $n = 0, \dots, N-1$, by construction the two terms inside the modulus converges to each other when Δt goes to infinity. Moreover, as in [9, Lemma 3.2] by Markovian property and central limit theorem, the rate of convergence has order $\frac{1}{\sqrt{\Delta t}}$ when Δt goes to infinity. Hence, for small ϵ we have

$$\begin{aligned}
\left[\mathbb{E} \sup_{s \leq t} |A_2|^p \right]^{\frac{1}{p}} &\leq K \epsilon N(\Delta t) \frac{1}{\sqrt{\Delta t}} \\
&\leq K \epsilon \left[\epsilon^{-1} |\ln \epsilon|^{-\frac{2\beta}{p}} \right] \frac{t |\ln \epsilon|^{\frac{2\beta}{p}}}{\sqrt{t} |\ln \epsilon|^{\frac{\beta}{p}}} \\
&= K \sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}}
\end{aligned}$$

□

Lemma 3.4 A_3 converges to zero when t or ϵ go to 0. Moreover, if Q^g is α -Hölder continuous with $0 < \delta < \alpha$ then the rate of convergence is given by

$$\left(\mathbb{E} \sup_{s \leq t} |A_3|^p \right)^{\frac{1}{p}} \leq K \epsilon t^{1+\frac{\delta}{2}} |\ln \epsilon|^{\frac{\beta\delta}{p}},$$

for a positive constant K .

Proof: By definition

$$|A_3| = \left| \sum_{n=0}^{N-1} \epsilon \Delta t Q^g(\pi(y_{t_n}^\epsilon)) - \epsilon \int_{\frac{s}{\epsilon} \wedge \tau^\epsilon}^{\frac{s+t}{\epsilon} \wedge \tau^\epsilon} Q^g(\pi(y_r^\epsilon)) dr \right|$$

and the convergence to zero here corresponds to the existence of the Riemann integral. Moreover, assuming that Q^g is α -Hölder continuous for $\delta < \alpha$; since y_t^ϵ is also α -Hölder continuous for $\alpha < \frac{1}{2}$ (classical result, see e.g. [11]), then we have that the previous expression is again bounded by

$$\begin{aligned} &\leq \epsilon \sum_{n=0}^{N-1} \Delta t \sup_{\epsilon t_n < s \leq \epsilon t_{n+1}} |Q^g((\pi(y_{t_n}^\epsilon))) - Q^g((\pi(y_s^\epsilon)))| \\ &\leq K \epsilon (\Delta t)^{(1+\frac{\delta}{2})} N \\ &\leq K \epsilon t^{1+\frac{\delta}{2}} |\ln \epsilon|^{\frac{\beta\delta}{p}}. \end{aligned}$$

□

Lemma 3.5 Process A_4 converges to zero with

$$\left(\mathbb{E} \sup_{s \leq t} |A_4|^p \right)^{\frac{1}{p}} \leq C t \epsilon |\ln \epsilon|^{\frac{2\beta}{p}}.$$

Proof: Denoting

$$C = \sup_{x \in U} |g(x)|.$$

The result follows straightforward since

$$\epsilon \left| \int_{t_N}^{\frac{s+t}{\epsilon} \wedge \tau^\epsilon} g(y_r^\epsilon) dr \right| \leq C \epsilon \Delta t = C t \epsilon |\ln \epsilon|^{\frac{2\beta}{p}}.$$

Now, going back to the proof of Lemma 3.1. Note that each of the four estimates of Lemmas 3.2–3.5 allows a factorization which has a common factor $\sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}}$ times a continuous function which goes to zero when $(t, \epsilon) \rightarrow 0$ (indeed, in Lemma 3.3, use a $\beta' \in (\beta, 1/2)$). Lemma 3.1 now follows by inequality (15).

□

4 An averaging principle

We state the averaging principle in the next theorem. To use Lemma 3.1 of the previous section we have to assume regularity in the average function Q^g , which naturally depends on g , on the foliated coordinate system and on the transversal behaviour of the invariant measures on the leaves of the original foliated system. We are going to assume the following:

Hypothesis (H): For any Lipschitz continuous function g on M , its corresponding average function Q^g on the transversal space V which indexes the leaves is also Lipschitz.

This hypothesis holds naturally if the invariant measures μ_p for the unperturbed foliated system has sort of weakly continuity on p . For deterministic systems it corresponds to a certain regularity in the sense that there is no bifurcation with respect to the vertical parameter $v \in V$.

We use the derivative of each component of

$$\pi(\cdot) = (\pi_1(\cdot), \dots, \pi_d(\cdot)) \in V \subset \mathbf{R}^d.$$

to get the averages $Q^{d\pi_i(K)}(x)$ of the real functions $d\pi_i(K)$, $i = 1, \dots, d$ on each leaf L_x .

Theorem 4.1 *Assume that the unperturbed foliated system (1) on M satisfies hypothesis (H) above. Let $v(t)$ be the solution of the deterministic ODE in the transversal component $V \subset \mathbf{R}^n$*

$$\frac{dv}{dt} = (Q^{d\pi_1(K)}, \dots, Q^{d\pi_d(K)})(v(t)) \quad (17)$$

with initial condition $v(0) = \pi(x_0) = 0$. Let T_0 be the time that $v(t)$ reaches the boundary of V , then

(1) *For all $0 < t < T_0$, $\beta \in (0, 1/2)$ and $2 \leq p < \infty$ ($1 \leq p$ for small t , cf. Remark 2.4),*

$$\left[\mathbb{E} \left(\sup_{s \leq t} \left| \pi \left(y_{\frac{s \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) - v(s) \right|^p \right) \right]^{\frac{1}{p}} \leq \sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}} h(t, \epsilon),$$

where $h(t, \epsilon)$ is continuous and converges to zero when ϵ or t goes to 0.

(2) *For $\gamma \leq 0$, let*

$$T_\gamma = \inf_{0 < t} \{ \text{dist}(v(t), \partial V) \leq \gamma \}$$

The exit times of the two systems satisfies the estimates

$$\mathbb{P}(\epsilon \tau^\epsilon < T_\gamma) \leq \gamma^{-p} t^{\frac{p}{2}} |\ln \epsilon|^{-\beta} h(t, \epsilon)^p.$$

Proof: The gradient of each real function π_i is orthogonal to the leaves, hence by Itô formula, for $i = 1, 2, \dots, d$ we have that

$$\pi_i \left(y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) = \int_0^{t \wedge \tau^\epsilon} d\pi_i(K)(y_{\frac{s}{\epsilon}}^\epsilon) ds.$$

Lemma 3.1 for the function $d\pi_i(K)$ in M , triangular inequality and hypothesis (H) imply that

$$\begin{aligned} \left| \pi_i \left(y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) - v_i(t) \right| &\leq \int_0^{t \wedge T^\epsilon} \left| Q^{d\pi_i(K)} \left(\pi \left(y_{\frac{s \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) \right) - Q^{d\pi_i(K)}(v(s)) \right| ds + |\delta_i(\epsilon, t)| \\ &\leq C_i \int_0^t \left| \pi \left(y_{\frac{s \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) - v(s) \right| ds + |\delta_i(\epsilon, t)|, \end{aligned}$$

where each C_i is the Lipschitz constant of $Q^{d\pi_i(K)}$ and $\delta_i(\epsilon, t)$ is defined in Lemma 3.1. Summing up the i 's and using Gronwall lemma we have, for a constant C :

$$\left| \pi \left(y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) - v(t) \right| \leq e^{Ct} \sum_{i=1}^n |\delta_i(\epsilon, t)|.$$

The first part of the theorem follows by Lemma 3.1. For the second part we have the following estimates

$$\begin{aligned} \mathbb{P}(\epsilon \tau^\epsilon < T_\gamma) &\leq \mathbb{P} \left(\sup_{s \leq T_\gamma \wedge \epsilon \tau^\epsilon} \left| v(s) - \pi \left(y_{\frac{s \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) \right| > \gamma \right) \\ &\leq \gamma^{-p} \mathbb{E} \left(\sup_{s \leq T_\gamma \wedge \epsilon \tau^\epsilon} \left| v(s) - \pi \left(y_{\frac{s \wedge \tau^\epsilon}{\epsilon}}^\epsilon \right) \right|^p \right) \\ &\leq \gamma^{-p} \left(\sqrt{t} |\ln \epsilon|^{-\frac{\beta}{p}} h(t, \epsilon) \right)^p \\ &\leq \gamma^{-p} t^{\frac{\beta}{2}} |\ln \epsilon|^{-\beta} h(t, \epsilon)^p. \end{aligned}$$

□

4.1 A detailed example:

The following simple example illustrates the framework where the averaging principle described in this section holds. Consider $M = \mathbf{R}^3 - \{(0, 0, z), z \in R\}$ with the 1-dimension horizontal circular foliation of M where the leaf passing through a point $p = (x, y, z)$ is given by the circle $L_p = \{(\sqrt{x^2 + y^2} \cos \theta, \sqrt{x^2 + y^2} \sin \theta, z), \theta \in [0, 2\pi]\}$. Consider the foliated linear SDE on M consisting of random rotations:

$$dx_t = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_t (\lambda_1 dt + \lambda_2 dB_t).$$

For an initial condition $p_0 = (x_0, y_0, z_0)$, say with $x_0 \geq 0$ consider the local foliated coordinates in the neighbourhood $U = \mathbf{R}^3 \setminus \{(x, 0, z); x \leq 0; z \in \mathbf{R}\}$ given

by cylindrical coordinates. Hence, using the same notation as before $\varphi = (u, v)$ will be defined by $\varphi : U \subset M \rightarrow (-\pi, \pi) \times \mathbf{R}_{>0} \times \mathbf{R}$, where $u \in (-\pi, \pi)$ is angular and $v = (r, z) \in \mathbf{R}_{>0} \times \mathbf{R}$ such that $\varphi^{-1} : (u, v) \mapsto (r \cos u, r \sin u, z) \in M$. In this coordinates system, the transversal projections π_1 and π_2 correspond to the radial r -component and the z -coordinate, respectively.

For $\lambda_1, \lambda_2 \in \mathbf{R}$ with $|\lambda_1| + |\lambda_2| > 0$, the invariant measures μ_p in the leaves L_p passing through points $p \in M$ are given by normalized Lebesgue measures in L_p , which here corresponds to the normalized angle 1-form. Note that Hypothesis (H) is satisfied. We investigate the effective behaviour of a small transversal perturbation of order ϵ :

$$dy_t^\epsilon = \lambda_1 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} y_t^\epsilon (\lambda_1 dt + \lambda_2 dB_t) + \epsilon K(y_t^\epsilon) dt.$$

with initial condition $x_0 = (1, 0, 0)$. In this example we shall consider two classes of perturbing vector field K .

Constant perturbation. (A) Assume that the perturbation is given by a vector field which is constant $K = (k_1, k_2, k_3)$ with respect to Euclidean coordinates in M . Initially, to fix the ideas, assume that $k_3 = 0$. Then, not only the average on the z -component vanishes, i.e. $Q^{d\pi_2 K} = 0$, but also, by the geometrical symmetry of K with respect to the invariant measure, the average radial r -component also vanishes, i.e. $Q^{d\pi_1 K} = 0$. Hence the transversal component in the main Theorem 4.1 is constant $v(t) = (r(0), z(0))$ for all $t \geq 0$.

Theorem 4.1 establishes a rate of convergence to zero of the difference between the initial radius $r(0) = 1$ and $r(\frac{t \wedge \tau^\epsilon}{\epsilon}) = \pi_1(y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon)$, the radial component of the perturbed systems, precisely, that

$$\left[\mathbb{E} \left(\sup_{s \leq t} \left| r\left(\frac{s \wedge \tau^\epsilon}{\epsilon}\right) - 1 \right|^p \right) \right]^{\frac{1}{p}}$$

goes to zero as ϵ and t goes to zero with the prescribed rate of convergence.

For comparison, indeed, the perturbed systems actually has solution

$$y_{\frac{t}{\epsilon}}^\epsilon = \begin{pmatrix} \cos\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) \\ \sin\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) \\ 0 \end{pmatrix} + \epsilon \begin{pmatrix} k_1 \sin\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) + k_2 \cos\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) - k_2 \\ -k_1 \cos\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) + k_2 \sin\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) - k_1 \\ 0 \end{pmatrix}$$

One can easily get an explicit description of the r -component: By normalization and using the symmetry, one can fix any k_1 and k_2 ; for simplicity, we shall fix $K = (1, 0, 0)$, hence, in this case, for $t \leq \tau^\epsilon$

$$r(t) = 1 + \epsilon^2 \left[2 + \cos\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) \right] - \epsilon \left[\cos 2\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) + \sin\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_{\frac{t}{\epsilon}}\right) \right]$$

Now, using that $r(t) + 1$ is bounded and that $|r(t) - 1| = |r^2(t) - 1|/|r(t) + 1|$ one finally finds that

$$|r(t) - 1| < K(t)\sqrt{\epsilon}$$

for any $t \leq \tau^\epsilon$, according to the boundaries on the the rate of convergence stated in the theorem.

(B) Vertical perturbations. Now, for constant and vertical $K = (0, 0, k_3)$, the radial average $Q^{d\pi_1 K}$ is null but $Q^{d\pi_2 K}$ equals k_3 for every leaf in M . Hence the averaged system $v(t) = (r(0), k_3 t)$ is constant in the radial component and increases linearly in the z -coordinate. The perturbed systems has the simple solution

$$y_t^\epsilon = \begin{pmatrix} \cos\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_t\right) \\ \sin\left(\frac{\lambda_1 t}{\epsilon} + \lambda_2 B_t\right) \\ k_3 t \end{pmatrix}$$

Hence, the comparison

$$|\pi_2\left(y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon\right) - v(t)| \equiv 0$$

for all $t \geq 0$ and the convergence of the theorem is trivially verified.

Linear perturbation. Consider a linear perturbation of the form $K(x, y, z) = (x, 0, 0)$. In this case, again, the z -coordinate average vanishes trivially. For the radial component, we have that $d\pi_1 K = r_0 \cos^2 u$, where u is the angular coordinate of p whose distance to the z -axis (r -coordinate) is r_0 . Hence the average with respect to the invariant measure on the leaves is given by $Q^{d\pi_1 K} = r/2$ for leaves with radius r . The transversal system stated in the Theorem is then $v(t) = (e^{\frac{t}{2}} r(0), z(0))$. Hence the result guarantees that the radial part of $y_{\frac{t \wedge \tau^\epsilon}{\epsilon}}^\epsilon$ must have a behaviour close to the exponential $e^{\frac{t}{2}}$ in the sense that

$$\left[\mathbb{E} \left(\sup_{s \leq t} \left| r\left(\frac{t}{\epsilon}\right) - e^{\frac{t}{2}} \right|^p \right) \right]^{\frac{1}{p}}$$

goes to zero when ϵ goes to zero.

The fundamental solution of the linear perturbed Stratonovich systems y_t^ϵ is given by the exponential of the matrix for each fixed t

$$\begin{pmatrix} \epsilon t & -(\lambda_1 t + \lambda_2 B_t) & 0 \\ \lambda_1 t + \lambda_2 B_t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where the eigenvalues corresponding to the first two coordinates (horizontal plane) are

$$\lambda_{1,2} = \frac{-\epsilon t \pm \sqrt{\epsilon^2 t^2 - 4(\lambda_1 t + \lambda_2 B_t)^2}}{2}.$$

whose real part is given by $\epsilon t/2$ with probability increasing to 1 as ϵ goes to zero. This exponential rates coincides with that one above guaranteed by the Theorem 4.1.

Lyapunov exponents. For foliated manifolds embedded in \mathbf{R}^N , the symmetry of the perturbing vector field K with respect to the geometry of the leaves, hence also with respect to Lebesgue invariant measure, as presented in the case of constant K , has implied that the transversal average $Q^{d\pi K}$ vanishes. This phenomenon also appears in a couple of other examples where the leaves are not only diffeomorphic to each other, but also has this symmetry in the sense that the integration of a constant perturbation $K \in \mathbf{R}^N$ with respect to the Lebesgue measure is zero. To mention a couple of simple examples: the spherical foliation of $\mathbf{R}^n \setminus \{0\}$, nested torus (increasing the smaller radius) foliation of the solid torus minus the central circle $S^1 \times D^2 \setminus S^1 \subset \mathbf{R}^3$ or more generally (when they exists) tubular foliation of $\mathbf{R}^n \setminus \{C\}$, with C a compact set (this context also includes the Hamiltonian case with the Lyouville foliation of the symplectic structure of \mathbf{R}^{2n} as in [9]). In these symmetric geometrical configuration, if the invariant measure on the leaves are the Lebesgue measures (taking gradient Brownian motion on the leaves for instance, as in [6]) the averages of $Q^{d\pi}$ vanishes, hence our main theorem says that on the average, the trajectories of the perturbed system stay somehow close to the initial leaf.

Lyapunov exponent of the system in the direction of a tangent vector $v \in T_{x_0}M$ contains the long time behaviour of points close to x_0 in the direction of v , for details on the definition, properties, existence conditions, multiplicative ergodic theory, etc see e.g. among many others L. Arnold [2] and the references therein. In particular, under the symmetric geometrical circumstances above, if there exists the Lyapunov exponents of the perturbed system y_t^ϵ , Theorem 4.1 will imply that in transversal directions the Lyapunov exponent can not be too far from zero. This vanishing property must happen with multiplicity given by the codimension of the foliation, as in the examples of the paragraphs above, where the asymptotic relevant parameters (rotation number ([12], [1], [5] and Lyapunov exponents) do exist. In short:

Proposition 4.2 (Continuity of Lyapunov exponents) *Assume that the perturbed system y_t^ϵ does have Lyapunov exponents a.s. at the assigned initial condition. If the averaged perturbation in the leaves vanishes, i.e. $Q^{d\pi K} = 0$, then a number, given by the codimension of the foliation, of Lyapunov exponents in the spectrum goes to zero as ϵ goes to zero.*

Proof: In fact, Theorem 4.1 says that the perturbed system increases in the transversal coordinates with order $|\pi(y_t^\epsilon)| \sim \sqrt{\epsilon t} h(\epsilon t, \epsilon)$. Hence, any exponential behaviour, if it exists, must be in the direction of the leaf L_{x_0} . □

Acknowledgments: This article has been written while the second author is in a sabbatical visit to University of Humboldt. He would like to express his gratitude

to Prof. Peter Imkeller and his research group for the nice and friendly hospitality. Author I. G.-G. has been supported by CNPq 555241/2009-2. and P. R. has been partially supported by FAPESP 11/50151-0 and 12/03992-1.

References

- [1] L. Arnold and P. Imkeller – Rotation numbers for linear stochastic differential equations. *Ann. Probab.* 27, 130–149 (1999).
- [2] L. Arnold – *Random Dynamical Systems*, Springer-Verlag (1998).
- [3] V. Arnold – *Mathematical Methods in Classical Mechanics*. 2nd ed. Berlin, Springer (1989).
- [4] A. Borodin and M. Friedlin – Fast oscillating random perturbations of dynamical systems with conservation laws. *Ann. Inst. H. Poincaré. Prob. Statist.*, 31 (1995) 485-525.
- [5] P. Catuogno, D. Ledesma and P. Ruffino – Relative rotation number for stochastic systems: dynamical and topological applications. *Dynamical Systems*, 23 (2008) 425-435.
- [6] P. Catuogno, D. Ledesma and P. Ruffino – Harmonic measures in embedded foliated manifolds. Submitted. (2012) ArXiv 1208.0629.
- [7] Y. Kabanov and S. Pergamenshchikov – *Two-Scale Stochastic Systems: asymptotic analysis and control*. Springer-Verlag, 2003.
- [8] R. Khasminski and N. Krylov – On averaging principle for diffusion processes with null-recurrent fast component. *Stoch Proc Appl.* 93 (2001) 229-240.
- [9] Xue-Mei Li – An averaging principle for a completely integrable stochastic Hamiltonian systems. *Nonlinearity*, 21 (2008) 803-822.
- [10] S. Namachchvaya and R. Sowers – Rigorous stochastic averaging at a center with additive noise. *Meccanica*, 37 (2002) 85-114.
- [11] D. Revuz and M. Yor – *Brownian Motion and Continuous Martingales*. Springer-Verlag, 3rd ed. Berlin 1999.
- [12] P. R. Ruffino – Decomposition of stochastic flow and rotation matrix. *Stochastics and Dynamics*, 2 (2002) 93-108.
- [13] J.A. Sanders, F. Verhulst and J. Murdock – *Averaging Methods in Nonlinear dynamical Systems*. 2nd Ed. Springer, 2007.
- [14] R. Sowers – Stochastic averaging with a flattened Hamiltonian: a Markov process on a stratified space (a whiskered sphere). *Trans. Am. Math. Soc.*, 354 (2002) 853-900.

- [15] P. Tondeur. *Foliations on Riemannian manifolds*. Universitext, Springer Verlag, Berlin-Heidelberg-New York, 1988.
- [16] P. Walczak – *Dynamics of foliations, groups and pseudogroups*. Birkhäuser Verlag 2004.